Optimal Tax Base with Administrative Fixed Costs

Stéphane Gauthier

Noname manuscript No.
(will be inserted by the editor)

Optimal Tax Base with Administrative Fixed Costs

Stéphane Gauthier

September 10, 2012

Abstract This note characterizes the optimal base for commodity taxation in the presence of administrative fixed costs varying across goods. For low tax rates, the optimal base only comprises commodities whose discouragement index is greater than the ratio of their administrative costs to the tax they yield. An illustration with UK data shows that a category of goods should be taxed only if the revenue generated on this category is at least ten times greater than its administrative fixed cost. The cost imputable to the category of goods taxed at the standard rate would be at most 6 percent of total VAT revenue. The administration cost associated with categories of goods currently tax free could justify exemption.

JEL classification numbers H21

Keywords indirect taxation · VAT · tax base · administrative costs

Stéphane Gauthier
Paris School of Economics and University of Paris 1

Stéphane Gauthier
106-112, bd de l'Hpital 75013 Paris (France)
Phone: +33144078289
E-mail: stephane.gauthier@univ-paris1.fr

1 Introduction

The theory of optimal taxation focuses on the excess burden of taxation as the main source of the social loss caused by taxation. The empirical evidence however suggests that the administrative costs required to collect taxes may be substantial (Slemrod (1991)). Most measures of such costs are based on staff salary or equipment costs, but little is known about their true underlying determinants and precise shape (Shaw et al (2010)). Polinsky and Shavell (1984), Kaplow (1990) and Mayshar (1991) have derived optimal tax rules when the administrative cost function displays usual continuity and convexity properties with respect to the level of tax rates. Still, as argued by Slemrod and Yitzhaki (1996) and Alm (1996), it is likely that this function exhibits significant discontinuities and/or nonconvexities.

Yitzhaki (1979) and Wilson (1989) have given early insights into the optimal indirect taxes when these discontinuities are due to idiosyncratic fixed costs varying across commodities. A minimum number of employees may be required for performing the administration of taxed goods. Otherwise, no tax would be recovered because of, e.g., tax evasion and black market operations. The tax authority then faces two alternatives: a category of goods can be either taxed, in which case the tax authority must bear some specific administration fixed cost, or this category can remain tax free. The characterization of the tax base therefore involves a discrete choice between taxation and exemption. In a simple Cobb-Douglas partial equilibrium economy with one representative consumer and uniform commodity taxation, Yitzhaki (1979) has shown that this choice relies on price sensitivity of demand, as is usual in a Ramsey setup, and also on the level of the demand. The tax base actually comprises all the goods with a low enough ratio of administrative cost to the amount of tax they yield. Hereafter this ratio will be referred to as the 'Yitzhaki ratio.'

This note generalizes the papers of Yitzhaki (1979) and Wilson (1989) in three directions: (i) optimal tax rates possibly differ across goods; (ii) households have heterogeneous preferences with respect to consumption and labor; (iii) the social planner has a redistributive objective. This generalization is made possible by considering a new problem which does not directly refer to the discrete choice between taxation and exemption, but instead the continuous choice of the probability that a category of similar goods will be subject to taxation. The optimal choice of this probability can be treated by appealing to the usual Lagrangian method, which enables us to compare solutions where the category is either taxed or exempted.

This construction yields a general rule for including a good in the tax base. This rule takes a very simple form in the particular case of low tax rates: a good must be included in the tax base only if its associated Mirrlees (1976) discouragement index is higher than its Yitzhaki ratio. Thus the influence of equity on the decision to tax only transits through the optimal levels of tax rates: a good whose demand should not be strongly discouraged, because it is consumed by agents whose social value is high, is less likely to be taxed.

An empirical illustration on UK data suggests that a category of goods should be taxed only if the revenue generated on this category is at least ten times greater than its administrative fixed cost. Currently untaxed categories of goods may fail to satisfy this requirement, which could justify exemption. The illustration also gives upper bounds for the ratio of administrative fixed costs to total VAT revenue. The highest bounds apply to the categories 'Household Goods and Services', 'Petrol and Diesel', 'Tobacco' and 'Leisure Services'. The administrative cost imputable to the whole category of goods taxed at the standard rate would be at most 6 per cent of total VAT revenue.

2 The setup

The economy consists of a tax authority, a continuum of heterogeneous households $h \in \mathcal{H}$, and a continuum of categories of consumption goods $i \in \mathcal{C}$ produced from labor provided by households. The tax authority raises a given amount of revenue R by using linear taxes on consumption goods. In the Ramsey approach the tax authority first sets the tax rates t_i applying to goods in category i, and then households choose how much to consume and how much to work, considering as given the various tax rates.

We proceed backward and begin with the study of households' behavior. This behavior will act as a constraint in the problem of the tax authority. The preferences of household h are represented by a utility function which is separable across consumption goods and labor,

$$\int_{\mathcal{C}} u_i^h(x_i^h)\mathrm{d}\mu_i - \ell^h, \tag{1}$$

where μ is a non-atomic measure which captures the importance of the different categories of goods.

The problem of household h consists in choosing a bundle (x_i^h) and a labor supply ℓ^h which maximize her utility (1) subject to the budget constraint

$$\int_{\mathcal{C}} (1+t_i)x_i^h\mathrm{d}\mu_i \leq \ell^h. \tag{2}$$

Consumption goods are produced from labor according to a linear technology normalized so that producer prices equal 1, and thus $1+t_i$ is the consumer price of goods in category i. The demand function solution to the problem of maximizing (1) subject to (2) is $x_i^h = \xi_i^h(t_i)$ for every $i \in \mathcal{C}$. Indirect utility is

$$\int_{\mathcal{C}} \left(u_i^h(\xi_i^h(t_i)) - (1+t_i)\xi_i^h(t_i)\right)\mathrm{d}\mu_i \equiv \int_{\mathcal{C}} v_i^h(t_i)\mathrm{d}\mu_i.$$

The contribution of category i goods to the welfare of household h can therefore be measured by $v_i^h(t_i)$.

Some of the arguments used in section 3 to characterize the composition of the tax base rely on the assumption that utility is derived from a continuum of goods (see Remark 1). The modeling assumption that goods differ continuously with some characteristics may not fully accord with the formulation (1), since a tax rate placed on good i is likely to cause substitution into almost identical neighboring goods. Section 4 considers a general utility function, with a continuum of consumption goods, allowing for cross-price and income effects. It shows that the simple formulation (1) actually provides the main insights.

3 The decision to tax

Following Yitzhaki (1979), the tax authority must pay a fixed cost c_i when it decides to tax (or subsidize) commodity i. Otherwise, this commodity remains tax free. The tax authority thus chooses both the tax rates applying to taxed commodities and the composition of the tax base.

Whether a commodity should be taxed or exempted is a discrete decision to which the standard Lagrangian method does not directly apply. It can be treated as a continuous decision by proceeding as if it were possible to tax commodities randomly. Suppose that commodity i is taxed at rate t_i with probability π_i ($0 \leq \pi_i \leq 1$) and exempted otherwise. When this commodity is taxed, household h gets indirect utility $v_i^h(t_i)$ and pays the tax $t_i \xi_i^h(t_i)$ while the authority bears the administration cost c_i. When it is tax free, household h indirect utility is $v_i^h(0)$.

The problem of the tax authority is to select t_i and π_i for every $i \in \mathcal{C}$. Let the measure ν stand for the distribution of households, and let γ^h be the social valuation of the welfare of household h. At the optimum, the profile (t_i, π_i) maximizes

$$\int_{\mathcal{H}} \gamma^h \left(\int_{\mathcal{C}} \pi_i v_i^h(t_i) \mathrm{d}\mu_i + \int_{\mathcal{C}} (1-\pi_i) v_i^h(0) \mathrm{d}\mu_i \right) \mathrm{d}\nu^h \tag{3}$$

subject to the budget constraint

$$\int_{\mathcal{C}} \pi_i \left(t_i \int_{\mathcal{H}} \xi_i^h(t_i) \, \mathrm{d}\nu^h - c_i \right) \mathrm{d}\mu_i \geq R \qquad (\lambda) \tag{4}$$

and for every $i \in \mathcal{C}$,

$$\pi_i \geq 0, \qquad (\rho_i)$$
$$\pi_i \leq 1. \qquad (\sigma_i)$$

The variables in brackets are the associated Lagrange multipliers.

Using the Lagrangian approach, a first-order (necessary) condition for π_i to be a maximum is

$$\int_{\mathcal{H}} \left(\gamma^h \left(v_i^h(t_i) - v_i^h(0) \right) + \lambda \left(t_i \xi_i^h(t_i) - c_i \right) \right) \mathrm{d}\nu^h + \rho_i - \sigma_i = 0. \tag{5}$$

In addition, the Kuhn and Tucker exclusion conditions must be satisfied,

$$\rho_i \geq 0, \quad \rho_i \pi_i = 0, \tag{6}$$

$$\sigma_i \geq 0, \quad \sigma_i(1-\pi_i) = 0. \tag{7}$$

Let

$$\mathcal{L}_i(t_i,\lambda) \equiv \int_{\mathcal{H}} \gamma^h v_i^h(t_i)\mathrm{d}\nu^h + \lambda t_i \int_{\mathcal{H}} \xi_i^h(t_i)\mathrm{d}\nu^h$$

stand for the contribution of commodity i to social welfare (net of its associated administrative costs). The first-order conditions (5), (6) and (7) directly yield the following result:

Proposition 1 *Assume that commodity i is taxed (or subsidized) at some arbitrary rate t_i when it belongs to the tax base. A necessary condition for commodity i to belong to the tax base is $\mathcal{L}_i(t_i,\lambda) - \mathcal{L}_i(0,\lambda) \geq \lambda c_i$, or equivalently*

$$\int_{\mathcal{H}} \beta^h \left(v_i^h(t_i) - v_i^h(0)\right) \mathrm{d}\nu^h + t_i \int_{\mathcal{H}} \xi_i^h(t_i)\mathrm{d}\nu^h \geq c_i, \tag{8}$$

where $\beta^h \equiv \gamma^h/\lambda$ is the marginal social valuation of the income of household h.

Similarly, a necessary condition for commodity i to be tax free is $\mathcal{L}_i(t_i,\lambda) - \mathcal{L}_i(0,\lambda) \leq \lambda c_i$, or equivalently

$$\int_{\mathcal{H}} \beta^h \left(v_i^h(t_i) - v_i^h(0)\right) \mathrm{d}\nu^h + t_i \int_{\mathcal{H}} \xi_i^h(t_i)\mathrm{d}\nu^h \leq c_i. \tag{9}$$

In the second stage described in Section 2 households will be faced with the situation where commodity i is either exempted ($\pi_i = 0$) or taxed ($\pi_i = 1$). The relevant solutions must therefore be such that $\pi_i \in \{0,1\}$. There is in fact no loss in focusing attention on this type of solutions in Proposition 1. This is clear when inequality (8) or (9) is strict. This is also clear when there is only one non-atomic commodity for which (8) holds at equality, since setting the probability on that commodity to either 0 or 1 can then be done without any further implications.

In the remaining case where there is an atomic group of commodities for which (8) holds at equality, and $\pi_i \notin \{0,1\}$ for some of these commodities, it is always possible to split this group into two (atomic) subgroups, one consisting of taxed commodities and the other consisting of exempted commodities, such that both the budget constraint and the social objective are unaffected. To see this, suppose that (8) holds at equality for every commodity in an atomic group $\mathcal{I} \subseteq \mathcal{C}$, i.e.,

$$\int_{\mathcal{H}} \gamma^h \left(v_i^h(t_i) - v_i^h(0)\right) \mathrm{d}\nu^h = -\lambda \left(t_i \int_{\mathcal{H}} \xi_i^h(t_i)\mathrm{d}\nu^h - c_i \right) \tag{10}$$

for all $i \in \mathcal{I}$, $\mu(\mathcal{I}) > 0$. Consider any interval $I = \left[i^{\inf}, i^{\sup}\right] \subseteq \mathcal{I}$, and suppose that $\pi_i \neq \{0,1\}$ for some $i \in I$. Let R_I be the amount collected from I. There is then a new profile (π_i^*), $\pi_i^* \in \{0,1\}$ for every $i \in I$, such that the profiles (π_i) and (π_i^*) yield the same amount R_I of collected tax,

$$\int_{\mathcal{I}} \pi_i \left(t_i \int_{\mathcal{H}} \xi_i^h \left(t_i\right) \mathrm{d}\nu^h - c_i \right) \mathrm{d}\mu_i = \int_{\mathcal{I}} \pi_i^* \left(t_i \int_{\mathcal{H}} \xi_i^h \left(t_i\right) \mathrm{d}\nu^h - c_i \right) \mathrm{d}\mu_i. \quad (11)$$

Indeed, the function

$$G(\underline{i}) = \int_{\underline{i}}^{i^{\sup}} \left(t_i \int_{\mathcal{H}} \xi_i^h \left(t_i\right) \mathrm{d}\nu^h - c_i \right) \mathrm{d}\mu_i$$

is continuous with $\underline{i}$, $\underline{i} \in \left[i^{\inf}, i^{\sup}\right]$, when the measure μ is non-atomic. Since $G(i^{\inf}) \geq R_I$ and $G(i^{\sup}) = 0$, it follows from the intermediate value theorem that there is at least one $\underline{i} \in I$ such that $G(\underline{i}) = R_I$. By (10), $G(i)$ is decreasing, so that such a $\underline{i}$ is in fact unique. Proceeding similarly for every interval in $\mathcal{I}$ yields a profile (π_i^*), $\pi_i^* \in \{0,1\}$ for every $i \in \mathcal{I}$, allowing to collect the same amount of tax from $\mathcal{I}$ as the initial profile (π_i). Finally, using (10), (11), and again (10) shows that the value of the social objective is also the same with both profiles,

$$\begin{aligned} &\int_{\mathcal{I}} \left(\int_{\mathcal{H}} \gamma^h v_i^h(0) \mathrm{d}\nu^h + \pi_i^* \int_{\mathcal{H}} \gamma^h \left(v_i^h(t_i) - v_i^h(0) \right) \mathrm{d}\nu^h \right) \mathrm{d}\mu_i \\ = &\int_{\mathcal{I}} \left(\int_{\mathcal{H}} \gamma^h v_i^h(0) \mathrm{d}\nu^h + \pi_i \int_{\mathcal{H}} \gamma^h \left(v_i^h(t_i) - v_i^h(0) \right) \mathrm{d}\nu^h \right) \mathrm{d}\mu_i. \end{aligned}$$

Remark 1 The existence of a solution where $\pi_i \in \{0,1\}$ for every good relies on the measure μ. When this measure is non-atomic, the previous argument shows that there may exist several solutions where $\pi_i \in [0,1]$ for every good, but there is always one solution where every good is either taxed or exempted. When the measure μ has atoms, e.g., in the discrete good version of this model, the function G used above is no longer continuous. Then it may not be possible to find a solution $\pi_i \in \{0,1\}$ for every good.

Proposition 1 characterizes the optimal tax base associated with arbitrary tax rates, not necessarily those chosen by the government to maximize social welfare. An optimal tax rate t_i^* on commodity i is an extremum of $\mathcal{L}_i(t_i, \lambda)$. It satisfies the first-order condition

$$-d_i^* \equiv \int_{\mathcal{H}} \frac{t_i^*}{\xi_i(t_i^*)} \frac{\partial \xi_i^h}{\partial t_i}(t_i^*) \mathrm{d}\nu^h = -\left((1-\beta) - \beta \phi_i^* \right), \quad (12)$$

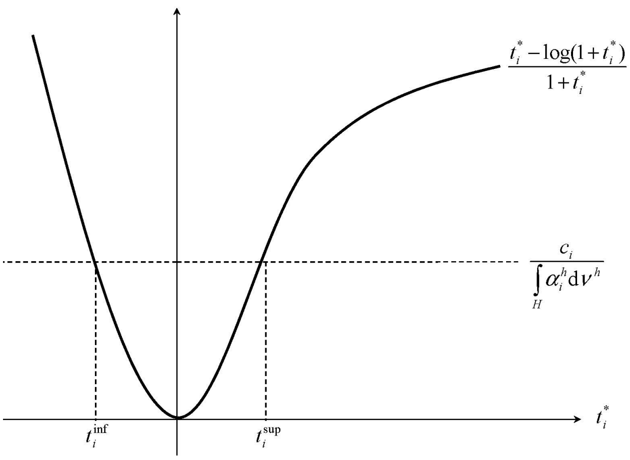

Fig. 1 The tax base in the Cobb-Douglas case

with

$$\beta \equiv \int_{\mathcal{H}} \beta^h \mathrm{d}\nu^h, \quad \xi_i(t_i^*) \equiv \int_{\mathcal{H}} \xi_i^h(t_i^*)\mathrm{d}\nu^h, \quad \phi_i^* \equiv \mathrm{cov}\left(\frac{\beta^h}{\beta}, \frac{\xi_i^h(t_i^*)}{\xi_i(t_i^*)}\right).$$

This is the familiar many-person Ramsey formula. At the optimum, the Mirrlees' (1976) discouragement index d_i^* equals a constant plus an equity correction varying with the commodity.

In an optimal indirect tax structure, the tax base comprises goods satisfying (8), the tax rates (t_i^*) satisfy (12) and the marginal social cost of public funds λ is determined by the budget constraint (4) of the tax authority.

In order to describe how the composition of the tax base described in Proposition 1 and the Ramsey rule (12) interact, consider for instance the particular case where preferences over consumption goods are represented by a Cobb-Douglas utility function, $u_i^h(x_i^h) = \alpha_i^h \log x_i^h$ for every h, i (whereas the contribution of leisure to utility remains a linear function of its consumption). Appealing to the Ramsey rule,

$$\int_{\mathcal{H}} \alpha_i^h \mathrm{d}\nu^h = (1+t_i^*) \int_{\mathcal{H}} \beta^h \alpha_i^h \mathrm{d}\nu^h,$$

inequality (8) becomes

$$\frac{t_i^* - \log(1+t_i^*)}{1+t_i^*} \geq \frac{c_i}{\int_{\mathcal{H}} \alpha_i^h \mathrm{d}\nu^h}. \tag{13}$$

In the representative agent case considered by Yitzhaki (1979), $\alpha_i^h = \alpha_i$ whatever h and i are, so that the Ramsey tax rate is uniform and the tax base consists of all the goods whose c_i/α_i ratio is low enough.

Figure 1 illustrates how the optimal tax base is determined in the presence of heterogeneity across households. The LHS of (13) is the U-shaped function of the tax rate, with a global minimum at $t_i^* = 0$. When satisfied at equality, (13) typically yields two thresholds $t_i^{\inf}$ and $t_i^{\sup}$ ($t_i^{\inf} < 0 < t_i^{\sup}$). Commodity i should be tax free only if $t_i^{\inf} \leq t_i^* \leq t_i^{\sup}$, and it should enter the tax base only if either $t_i^* \geq t_i^{\sup}$ or $t_i^* \leq t_i^{\inf}$.

Inequality (13) shows that, in the Cobb-Douglas case, heterogeneity across households and equity considerations influence the composition of the tax base through the Ramsey tax rates only. This property extends to more general preferences satisfying (1) in the empirically plausible configuration of low rates of tax, i.e., when t_i^* is close enough to 0 for every $i \in \mathcal{C}$.

Proposition 2 *Assume that commodity i is taxed at a low Ramsey tax rate t_i^* when it belongs to the tax base. It should be included in the tax base only if its associated discouragement index is greater than its Yitzhaki ratio, i.e.,*

$$|d_i^*| \geq \frac{c_i}{|t_i^*|\, \xi_i(0)}.$$

Proof Appealing to Roy's identity, the first-order Taylor expansion of $\mathcal{L}_i(t_i^*, \lambda)$ at $t_i^* = 0$ yields

$$\mathcal{L}_i(t_i^*, \lambda) \simeq \mathcal{L}_i(0, \lambda) + \lambda t_i^* \xi_i(0) \left(1 - \int\limits_{\mathcal{H}} \beta^h \frac{\xi_i^h(0)}{\xi_i(0)} \mathrm{d}\nu^h \right). \tag{14}$$

The tax rate t_i^* is given by (12), which can be rewritten as

$$d_i^* = 1 - \int\limits_{\mathcal{H}} \beta^h \frac{\xi_i^h(t_i^*)}{\xi_i(t_i^*)} \mathrm{d}\nu^h. \tag{15}$$

For low tax rates, $t_i^* \simeq 0$ for every $i \in \mathcal{C}$,

$$\int\limits_{\mathcal{H}} \beta^h \frac{\xi_i^h(t_i^*)}{\xi_i(t_i^*)} \mathrm{d}\nu^h \simeq \int\limits_{\mathcal{H}} \beta^h \frac{\xi_i^h(0)}{\xi_i(0)} \mathrm{d}\nu^h + t_i^* \int\limits_{\mathcal{H}} \frac{\partial}{\partial t_i^*} \left(\beta^h \frac{\xi_i^h(t_i^*)}{\xi_i(t_i^*)} \right) \Bigg|_{t_i^*=0} \mathrm{d}\nu^h,$$

so that (15) yields

$$1 - \int\limits_{\mathcal{H}} \beta^h \frac{\xi_i^h(0)}{\xi_i(0)} \mathrm{d}\nu^h \simeq d_i^* + t_i^* \int\limits_{\mathcal{H}} \frac{\partial}{\partial t_i^*} \left(\beta^h \frac{\xi_i^h(t_i^*)}{\xi_i(t_i^*)} \right) \Bigg|_{t_i^*=0} \mathrm{d}\nu^h.$$

Reintroducing this expression into (14), and neglecting the higher order term in t_i^{*2}, one gets

$$\mathcal{L}_i(t_i^*, \lambda) \simeq \mathcal{L}_i(0, \lambda) + \lambda t_i^* \xi_i(0) d_i^*. \tag{16}$$

The result then follows from Proposition 1.

Proposition 2 gives a simple picture of the optimal tax base. Assume for instance that the Ramsey tax rate t_i^* is positive. Then, commodity i is more likely to be exempted whenever (a) it is costly to administrate (c_i is high), (b) it yields a low amount of taxes ($t_i^* \xi_i(0)$ is low), and (c) its demand should not be strongly discouraged ($d_i^* > 0$ is low), i.e., the efficiency cost induced by taxation of commodity i is high (β is high) and this good is consumed by households with high social value (ϕ_i^* is positive).

4 An empirical illustration

The method used in Sections 2 and 3 can be adapted to characterize the tax base in the presence of cross-price and income effects. Let the preferences of household h be represented by a general utility function $u^h(x, \ell)$ with usual monotony and convexity properties. When this household is faced with tax rate $t_i = \{0, t_i^*\}$ on commodity i, her budget constraint is

$$\int_{\mathcal{C}} (1 + t_i) x_i^h \mathrm{d}\mu_i \leq \ell^h.$$

The demand function of household h for good i is $\xi_i^h(t)$ and her labor supply is $\ell^h(t)$, where $t = (t_i)$ denotes the whole vector of tax rates. Her indirect utility is $v^h(t)$.

As in Section 3 we assume that the tax rate t_i is set according to a Bernoulli distribution with parameter π_i $(0 \leq \pi_i \leq 1)$. Social welfare then can be expressed as

$$\int_{\mathcal{H}} \gamma^h \mathbb{E}_t \left[v^h(t) \right] \mathrm{d}\nu^h \tag{17}$$

and the social budget constraint is

$$\int_{\mathcal{H}} \mathbb{E}_t \left[\int_{\mathcal{C}} t_i \xi_i^h(t) \mathrm{d}\mu_i \right] \mathrm{d}\nu^h - \int_{\mathcal{C}} \pi_i c_i \mathrm{d}\mu_i \geq R. \tag{18}$$

The optimal indirect tax structure is a profile (t_i^*, π_i^*) which maximizes (17) subject to (18) and the constraints $0 \leq \pi_i \leq 1$ for every $i \in \mathcal{C}$. When the optimal tax rates are low enough, it is shown in appendix that commodity i should be included in the tax base, i.e., $\pi_i^* = 1$, only if

$$\left(1 - \int_{\mathcal{H}} \beta^h \frac{\xi_i^h(0)}{\xi_i(0)} \mathrm{d}\nu^h \right) t_i^* \simeq d_i^* t_i^* \geq \frac{c_i}{\xi_i(0)}, \tag{19}$$

where $\beta^h = \alpha^h \gamma^h / \lambda$, and α^h is the marginal utility of income of household h. This generalizes the inequality given in Proposition 2 for a utility function for which the derived demands are income sensitive and for which there are cross-price effects.

Inequality (19) can be used to get bounds for actual administrative costs. Appealing to UK data from the Institute of Fiscal Studies, goods are grouped in twenty homogeneous discrete categories of goods reported in Table 1. Approximating integration to summation, the necessary condition (19) will be applied to these categories of goods. All the goods in a category are taxed at the same rate (except for goods subject to excises). Belan et al (2008) recovered from these data a marginal social cost of public funds λ equal to 1.11, and a poorly redistributive VAT with $\alpha_1\gamma_1 = 0.03$, $\alpha_4\gamma_4 = 0.54$ and $\alpha_5\gamma_5 = 0.43$, while $\alpha_h\gamma_h = 0$ for all the other deciles of consumption expenditures. The budget shares by decile for these categories then enable us to compute the actual discouragement indices (applying to uncompensated demand) which appear in the LHS of inequality (19). They are reported in the second column of Table 1, with the shares of VAT from each category in the total revenue (in the third column).

The quantities in (19) are evaluated at $t = 0$ since the rule (19) has been obtained under the assumption that the tax rates are close to 0. On the other hand, the computation of the discouragement indices is made using the actual (observed) VAT structure, with some goods being taxed. This computation is therefore valid provided that the quantities in (19) are not too sensitive to changes in the tax rates. With this caveat in mind, each index is found positive and close to 0.1. Hence, by (19), a category will be taxed only if the tax it yields is at least ten times greater than its fixed administrative cost.

Existing empirical evidence suggests that total VAT administrative costs, i.e., including variable costs but excluding possible compliance costs, would be around 1 or 2 percent of total VAT revenue in the UK (Sandford et al (1989); Bickley (2003); Evans (2003)). In order to check whether these estimates are consistent with those obtained from (19), inequality (19) is expressed in share of total VAT revenue,

$$\frac{t_k^*\xi_k(0)}{\displaystyle\int_{\mathcal{C}} t_i^*\xi_i(0)\mathrm{d}\mu_i} d_k^* \geq \frac{c_k}{\displaystyle\int_{\mathcal{C}} t_i^*\xi_i(0)\mathrm{d}\mu_i} \tag{20}$$

for every category k. Assuming that the actual (observed) demand is given by $\xi_i(t_i^*)$ and can be evaluated as if the various tax rates were zero, the observed share of VAT from every taxed category in total VAT revenue appears in the LHS of inequality (20), while the RHS gives the actual ratio of administrative fixed costs to total VAT revenue. The product of the share of VAT from every taxed category in total revenue and the discouragement index of the category thus provides an upper bound for the ratio of administrative fixed costs to total VAT revenue. These upper bounds for the ratio of administrative fixed costs to total VAT revenue are reported in the fourth column of Table 1. Their highest values stand between 1% and 2% of total VAT revenue. They apply to four categories ('Tobacco', 'Petrol and Diesel', 'Household Goods and Services', and 'Leisure Services'). Otherwise the administrative fixed costs seem low, e.g., those relative to 'Public Transport' do not exceed 0.2% of total VAT revenue, that is

$$0.002 \geq \frac{c_k}{\int\limits_{\mathcal{C}} t_i^* \xi_i(0) \mathrm{d}\mu_i}$$

for $k =$ 'Public Transport.' An upper bound for the ratio of total administrative fixed cost to total VAT revenue

$$\int_{\mathcal{T}} \frac{c_k}{\int\limits_{\mathcal{C}} t_i^* \xi_i(0) \mathrm{d}\mu_i}$$

can be found by summation of the bounds for the individual taxed goods. The fixed cost imputable to the whole category of goods taxed at the standard rate would be at most 5% and 6% of total revenue. These upper bounds on administrative fixed costs are above the existing estimates of the total administrative costs, and they have approximately the same order of magnitude as these estimates. One may therefore conclude that the results derived from (19) are roughly consistent with the existing UK estimates.

Finally one can also get an idea about the costs attached to untaxed categories by considering what would happen if the optimal rate on these categories were the reduced rate of VAT of 5%. The fifth column of Table 1 reports the corresponding bounds for administrative fixed costs, postulating no reaction of demand. These bounds can be viewed as lower bounds for the ratio of administrative costs of untaxed categories to total VAT revenue. These bounds seem low, about 0.1 or 0.2% of total revenue. In comparison with the actual administrative costs found in the literature, it seems therefore possible that these categories should remain tax free, partly because of relatively high fixed administrative costs.

	t_k	d_k	$\dfrac{t_k \xi_k}{\int_{\mathcal{C}} t_i \xi_i \mathrm{d}\mu_i}$	$\dfrac{t_k \xi_k}{\int_{\mathcal{C}} t_i \xi_i \mathrm{d}\mu_i} d_k$	$\dfrac{\tilde{t}_k \xi_k}{\int_{\mathcal{C}} \tilde{t}_i \xi_i \mathrm{d}\mu_i} d_k$
Untaxed					
Meat & Fish	0	0.0899	0	–	0.0023
Bread & Cereals	0	0.0901	0	–	0.0007
Dairy	0	0.0900	0	–	0.0015
Tea & Coffee	0	0.0901	0	–	0.0004
Fruits and Vegetables	0	0.0902	0	–	0.0014
Other Untax. Foods	0	0.0903	0	–	0.001
Books & Newspapers	0	0.0905	0	–	0.0006
Children's Clothing	0	0.0907	0	–	0.0003
Reduced rate					
Domestic Fuels	0.05	0.0900	0.0356	0.0032	–
Excises					–
Tobacco	0.791	0.0890	0.1379	0.012	–
Beer	0.28	0.0904	0.0634	0.0057	–
Petrol and Diesel	0.732	0.0908	0.1465	0.0133	–
Wine & Spirits	0.558	0.0913	0.0538	0.0049	–
Standard rate					–
Public Transport	0.175	0.0903	0.0211	0.0019	–
Standard VAT Food	0.175	0.0904	0.0381	0.0034	–
Food out	0.175	0.0919	0.0590	0.0054	–
Household G & S	0.175	0.0912	0.2005	0.0182	–
Adult Clothing	0.175	0.0914	0.0806	0.0073	–
Leisure Goods	0.175	0.0926	0.0508	0.0047	–
Leisure Services	0.175	0.0931	0.1101	0.010	–

Table 1 UK VAT base with administrative fixed costs

Acknowledgements I thank Gaël Giraud, Guy Laroque, Bertrand Wigniolle, and Shlomo Yitzhaki for thoughtful comments on the paper. I also thank Eckhard Janeba, editor of this Journal, and two anonymous referees. I am particularly grateful to one reviewer for carefully reading my manuscript, providing guidance and constructive comments which helped me to improve the paper. The usual disclaimers apply.

References

Alm J (1996) What is an optimal tax system? National Tax Journal 49: 117–33
Belan P, Gauthier S, Laroque G (2008) Optimal grouping of commodities for indirect taxation. Journal of Public Economics 92(7): 1738 – 1750
Bickley J (2003) Value Added Tax: Concepts, Policy Issues, and OECD Experiences. Novinka Books, New-York
Evans C (2003) Studying the studies: An overview of recent research into taxation operating costs. eJournal of Tax Research 1
Kaplow L (1990) Optimal taxation with costly enforcement and evasion. Journal of Public Economics 43: 221–236
Mayshar J (1991) Taxation with costly administration. Scandinavian Journal of Economics 93: 75–88
Mirrlees J (1976) Optimal tax theory : A synthesis. Journal of Public Economics 6(4): 327–358
Polinsky AM, Shavell S (1984) Pigouvian taxation with administrative costs. NBER Working Papers 742, National Bureau of Economic Research
Sandford C, Godwin M, Hardwick P (1989) Administrative and Compliance Costs of Taxation. Fiscal Publications
Shaw J, Slemrod J, Whiting J (2010) Administration and compliance. URL http://www.ifs.org.uk/mirrleesreview/dimensions/ch12.pdf
Slemrod J (1991) Optimal taxation and optimal tax systems. NBER Working Papers 3038, National Bureau of Economic Research
Slemrod J, Yitzhaki S (1996) The costs of taxation and the marginal efficiency cost of funds. Staff Papers - International Monetary Fund 43(1): 172–198
Wilson JD (1989) On the optimal tax base for commodity taxation. American Economic Review 79(5): 1196–1206
Yitzhaki S (1979) A note on optimal taxation and administrative costs. The American Economic Review 69(3): 475–480

Appendix

This appendix provides an explicit derivation of inequality (19). Let $\mathcal{T}$ be the set of all taxed commodities, i.e., $t_i \neq 0$ for every $i \in \mathcal{T}$, and $t_i = 0$ otherwise ($i \in \mathcal{C} \backslash \mathcal{T}$). The optimal tax rate t_k^* on commodity k is an extremum of the Lagrangian function

$$\int_{\mathcal{H}} \gamma^h v^h \mathrm{d}\nu^h + \lambda \int_{\mathcal{H}} \int_{\mathcal{T}} \left(t_i \xi_i^h - c_i \right) \mathrm{d}\mu_i \mathrm{d}\nu^h - \lambda R,$$

where v^h and ξ_i^h are evaluated at (t). The first-order condition in t_k writes

$$d_k^* \equiv - \int_{\mathcal{H}} \int_{\mathcal{T}} \frac{t_i}{\xi_k} \frac{\partial \xi_i^h}{\partial t_k} \mathrm{d}\mu_i \mathrm{d}\nu^h = \int_{\mathcal{H}} \left(1 - \frac{\gamma^h \alpha^h}{\lambda} \right) \frac{\xi_k^h}{\xi_k} \mathrm{d}\nu^h$$

where ξ_k is the aggregate consumption of good k, α^h is the marginal utility of income for household h, and d_k^* is the discouragement index for category k. All these quantities are evaluated at (t^*). Note that, unlike the usual formulation, the discouragement indices refer to uncompensated demand (and not compensated demand) in order to be used in the empirical illustration. The optimal tax rate on good k satisfies

$$d_k^* = 1 - \int_{\mathcal{H}} \beta^h \frac{\xi_k^h(t)}{\xi_k(t)} \mathrm{d}\nu^h \simeq 1 - \int_{\mathcal{H}} \beta^h \frac{\xi_k^h(0)}{\xi_k(0)} \mathrm{d}\nu^h + \int_{\mathcal{T}} t_i^* \left. \frac{\partial d_k^*}{\partial t_i} \right|_{t_i^*=0} \mathrm{d}\mu_i, \qquad (21)$$

where the approximation is obtained by assuming low tax rates.

Let $t_i = \{0, t_i^*\}$ and assume that $t_i \sim t_i^* \times \mathcal{B}(\pi_i)$, where $\mathcal{B}(\pi_i)$ stands for the Bernoulli distribution with parameter π_i. In order to derive the first-order condition associated with the inclusion of commodity k in the tax base, we let $t = (t_k, t_{-k})$ where t_{-k} is the set of all the tax rates different from t_k. The social welfare becomes

$$\pi_k \mathbb{E}_{t_{-k}} \left(\int_{\mathcal{H}} \gamma^h v^h(t_k^*, t_{-k}) \mathrm{d}\nu^h \right) + (1 - \pi_k) \mathbb{E}_{t_{-k}} \left(\int_{\mathcal{H}} \gamma^h v^h(0, t_{-k}) \mathrm{d}\nu^h \right),$$

while the total receipt from commodity taxation is

$$\pi_k \mathbb{E}_{t_{-k}} \left(t_k^* \int_{\mathcal{H}} \xi_k^h(t_k^*, t_{-k}) \mathrm{d}\mu_k \mathrm{d}\nu^h + \int_{\mathcal{H}} \int_{\mathcal{C}\backslash\{k\}} t_i \xi_i^h(t_k^*, t_{-k}) \mathrm{d}\mu_i \mathrm{d}\nu^h \right) + (1 - \pi_k) \mathbb{E}_{t_{-k}} \left(\int_{\mathcal{H}} \int_{\mathcal{C}\backslash\{k\}} t_i \xi_i^h(0, t_{-k}) \mathrm{d}\mu_i \mathrm{d}\nu^h \right).$$

Let also

$$\mathcal{L}_k(t_k^*, \lambda) = \int_{\mathcal{H}} \gamma^h \mathbb{E}_{t_{-k}} \left[v^h(t_k^*, t_{-k}) \right] \mathrm{d}\nu^h + \lambda \mathbb{E}_{t_{-k}} \left(\int_{\mathcal{H}} t_k^* \xi_k^h(t_k^*, t_{-k}) \mathrm{d}\mu_k \mathrm{d}\nu^h + \int_{\mathcal{H}} \int_{\mathcal{C}\setminus\{k\}} t_i \xi_i^h(t_k^*, t_{-k}) \mathrm{d}\mu_i \mathrm{d}\nu^h \right).$$

By the same argument as the one used to get Proposition 1, commodity k should be included in the tax base only if

$$\mathcal{L}_k(t_k^*, \lambda) - \mathcal{L}_k(0, \lambda) - \lambda c_k \geq 0. \tag{22}$$

For low tax rates, i.e., t_i^* close to 0, Roy's identity yields

$$v^h(t_k^*, t_{-k}) \simeq v^h - \int_{\mathcal{T}} t_i^* \alpha^h \xi_i^h \mathrm{d}\mu_i$$

where quantities are evaluated at $(t_k^*, t_{-k}) = (0, 0)$. Therefore, at this point,

$$\mathcal{L}_k(t_k^*, \lambda) \simeq \int_{\mathcal{H}} \gamma^h \left(v^h - \int_{\mathcal{T}} t_i^* \alpha^h \xi_i^h \mathrm{d}\mu_i \right) \mathrm{d}\nu^h + \lambda \int_{\mathcal{H}} \int_{\mathcal{T}} t_i^* \xi_i^h \mathrm{d}\mu_i \mathrm{d}\nu^h.$$

Similarly,

$$\mathcal{L}_k(0, \lambda) \simeq \int_{\mathcal{H}} \gamma^h \left(v^h - \int_{\mathcal{T}\setminus\{k\}} t_i^* \alpha^h \xi_i^h \mathrm{d}\mu_i \right) \mathrm{d}\nu^h + \lambda \int_{\mathcal{H}} \int_{\mathcal{T}\setminus\{k\}} t_i^* \xi_i^h \mathrm{d}\mu_i \mathrm{d}\nu^h.$$

The last two expressions only differ with respect to the inclusion of commodity k in the tax base. This commodity belongs to the tax base $\mathcal{T}$ in $\mathcal{L}_k(t_k^*, \lambda)$, but it is not in $\mathcal{T}$ (in fact it belongs to $\mathcal{T}\setminus\{k\}$) in $\mathcal{L}_k(0, \lambda)$. It follows that

$$\mathcal{L}_k(t_k^*, \lambda) - \mathcal{L}_k(0, \lambda) \simeq \int_{\mathcal{H}} \left(\lambda - \gamma^h \alpha^h \right) t_k^* \xi_k^h \mathrm{d}\nu^h = \lambda t_k^* \xi_k \left(1 - \int_{\mathcal{H}} \beta^h \frac{\xi_k^h}{\xi_k} \mathrm{d}\nu^h \right).$$

Appealing to (21), this approximation rewrites

$$\mathcal{L}_k(t_k^*, \lambda) - \mathcal{L}_k(0, \lambda) \simeq \lambda t_k^* \xi_k \left(d_k^* - \int_{\mathcal{T}} t_i^* \frac{\partial d_k^*}{\partial t_i} \mathrm{d}\mu_i \right) \simeq \lambda t_k^* \xi_k d_k^* \tag{23}$$

where the last approximation is obtained by neglecting the higher order terms in $t_k^* t_i^*$. Inequality (19) then comes from (22) and (23).